

Figure 1 – Main Concept

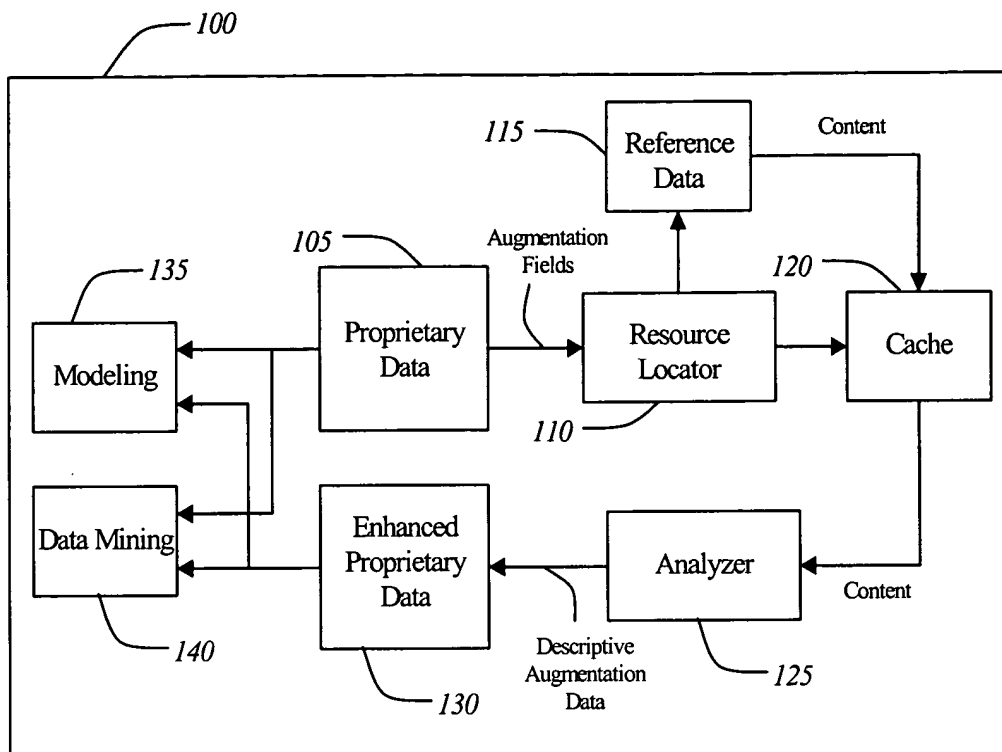


Figure 2 – Credit Card Example

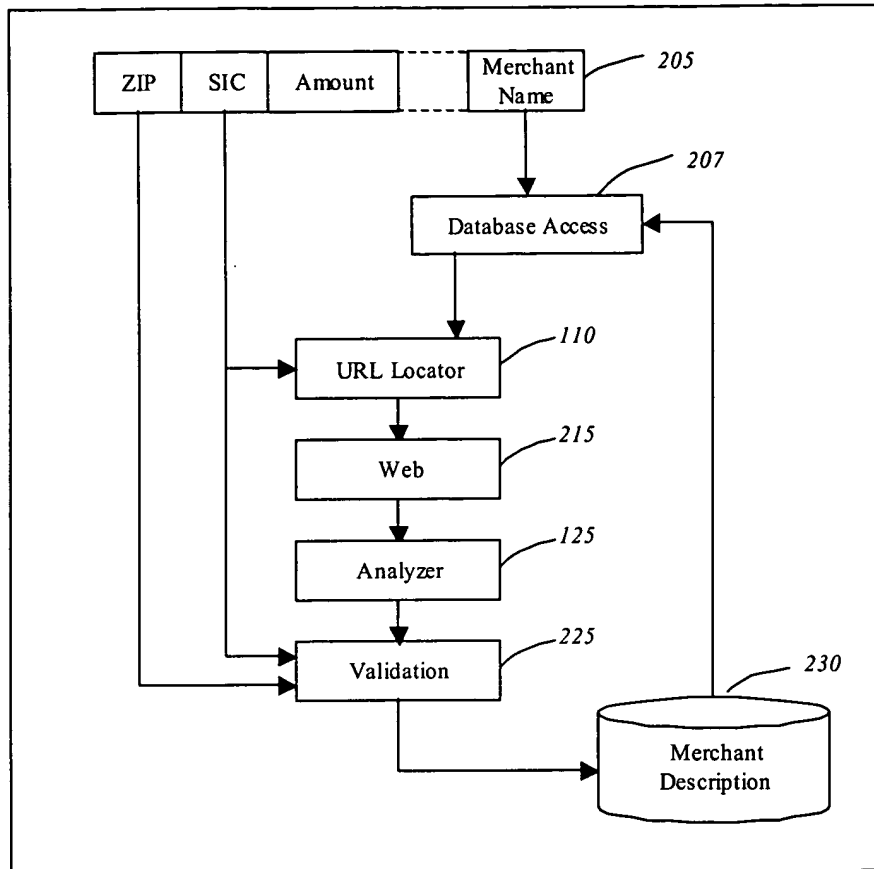
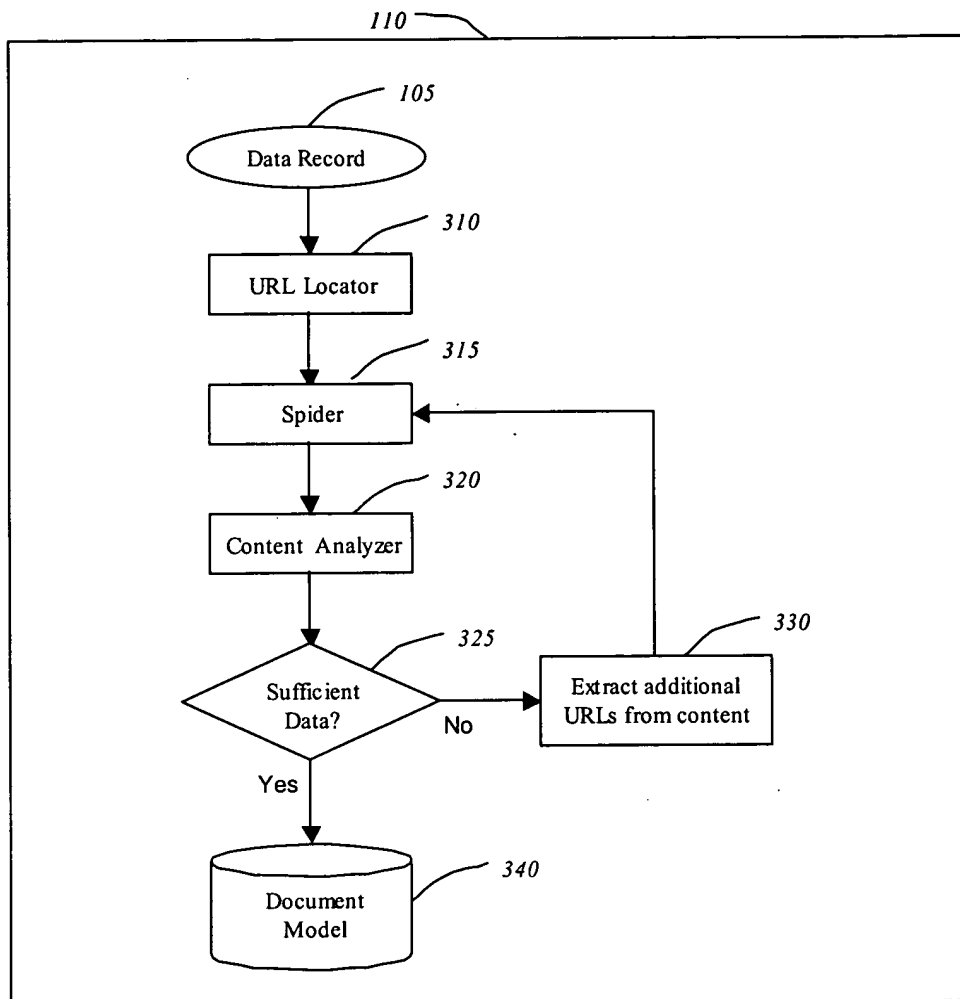


Figure 3 – Web Locator



100322643 "122701

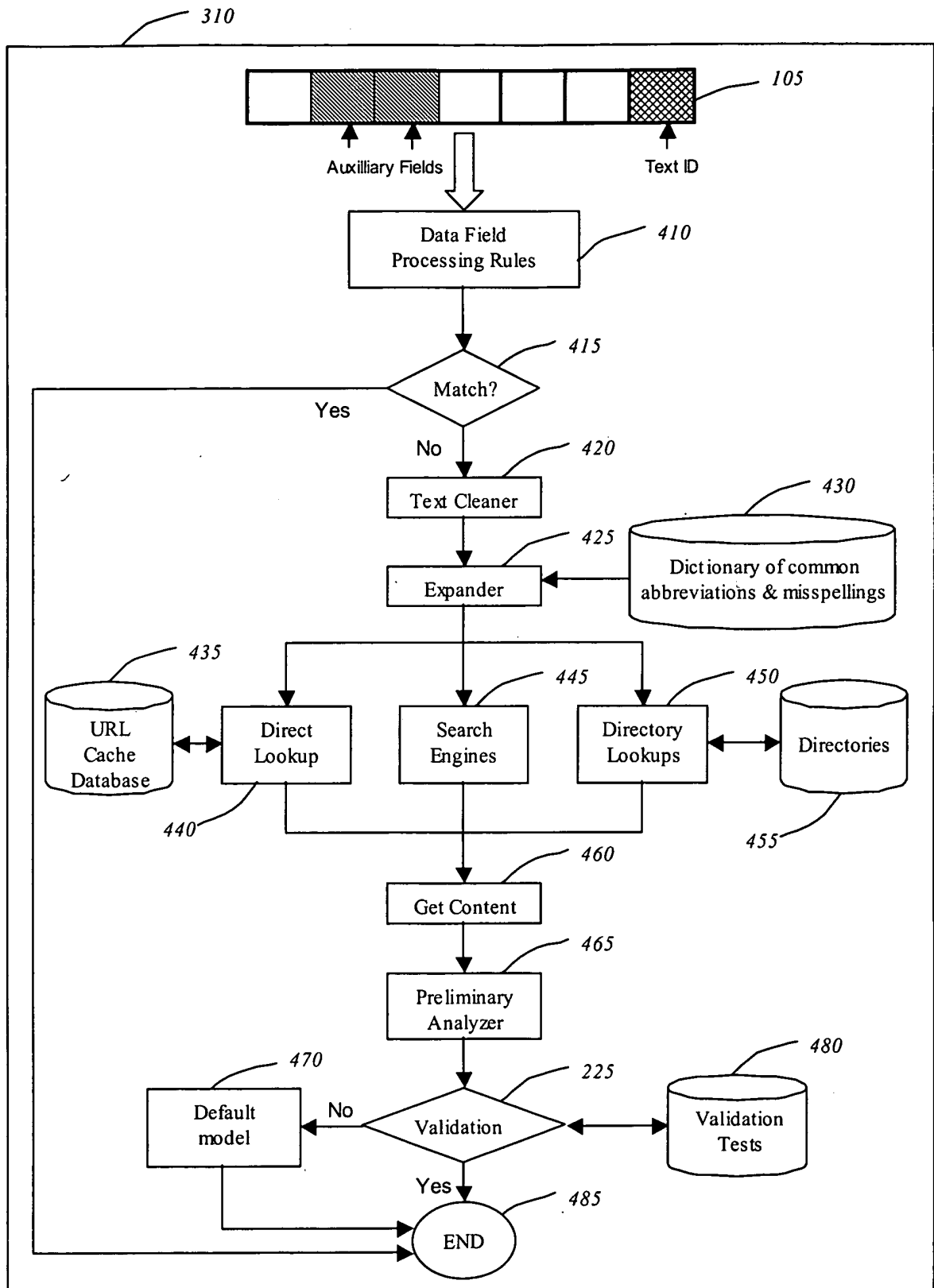
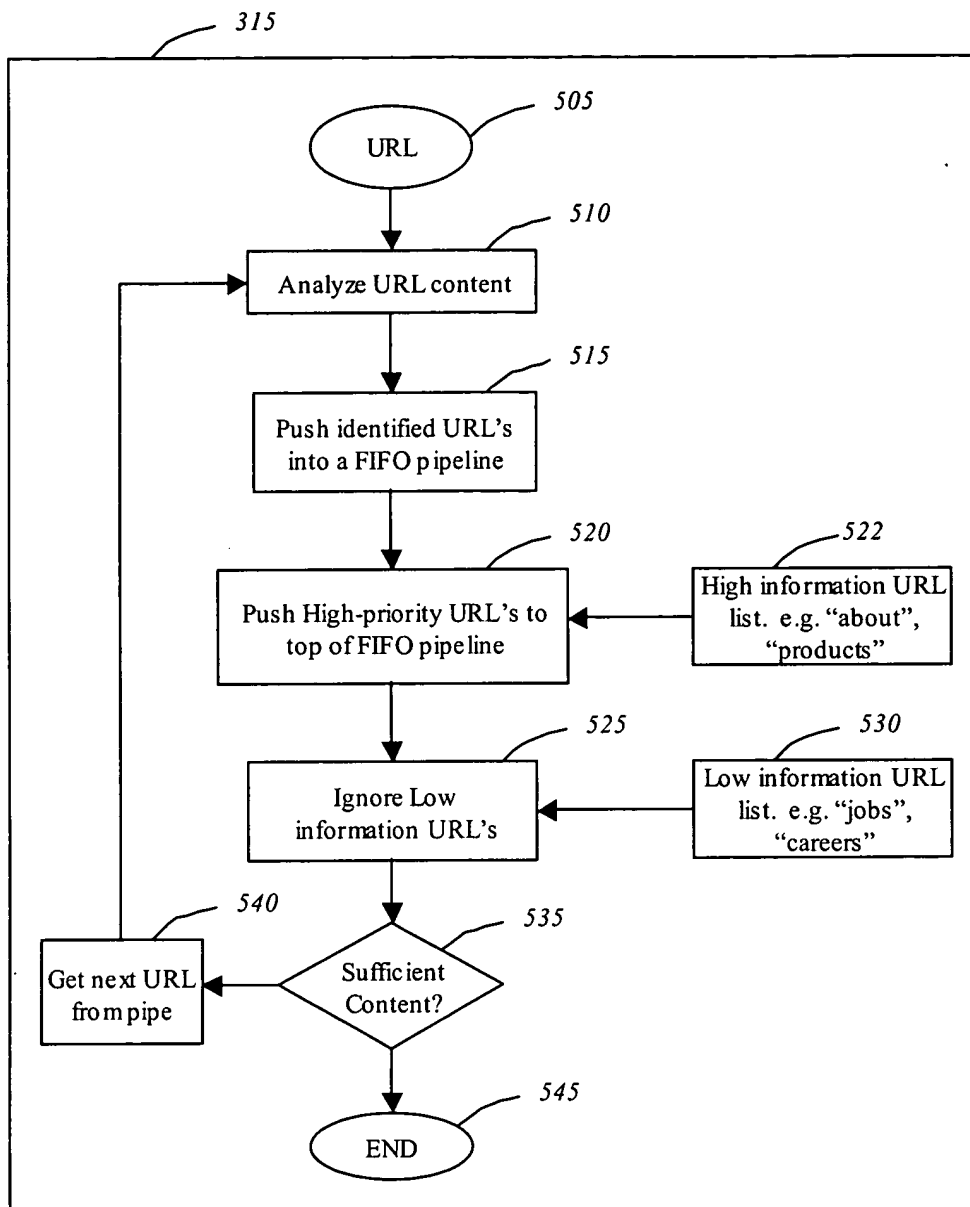


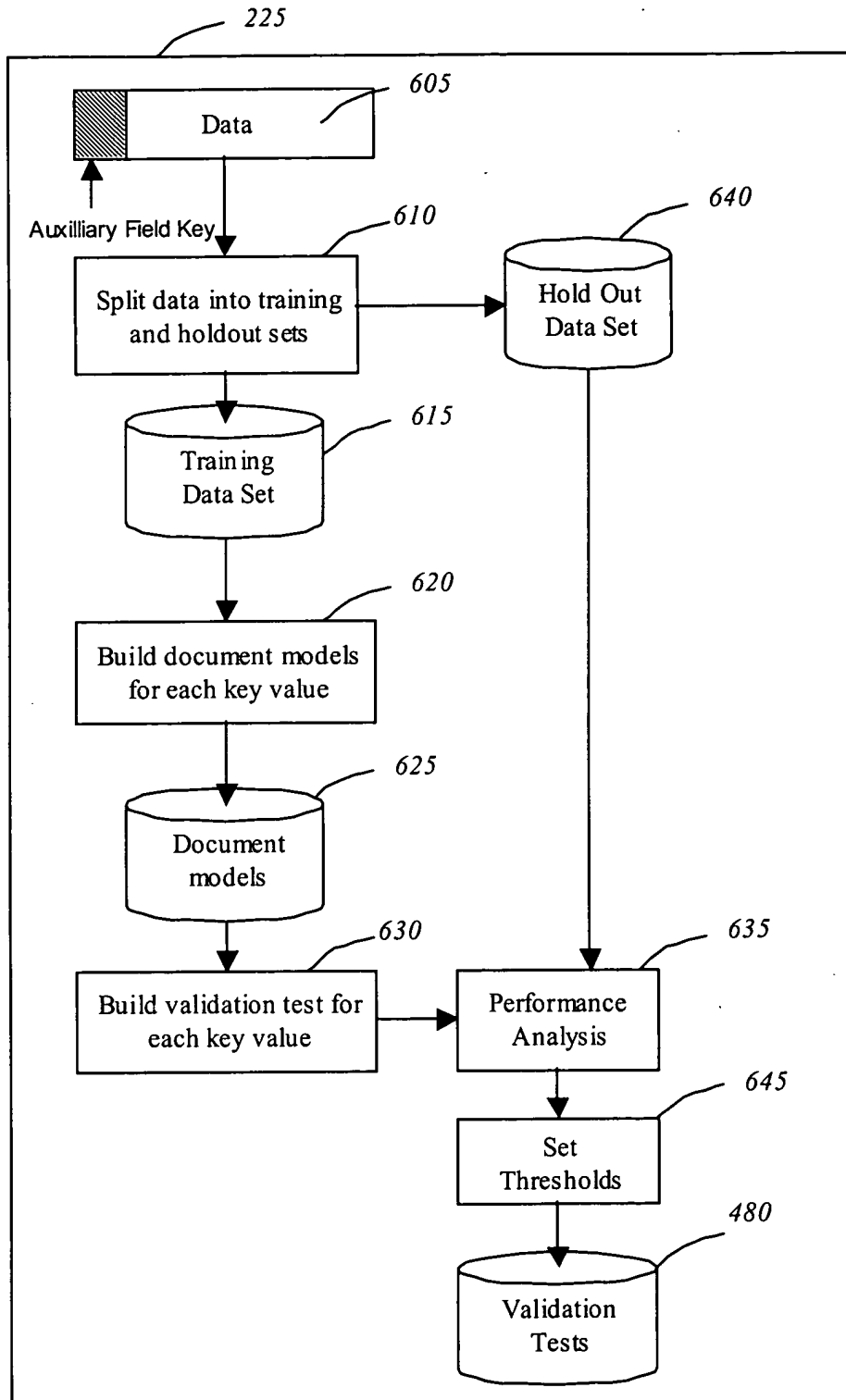
Figure 4 – URL Locator Pipeline

Figure 5 – Spidering Pipeline



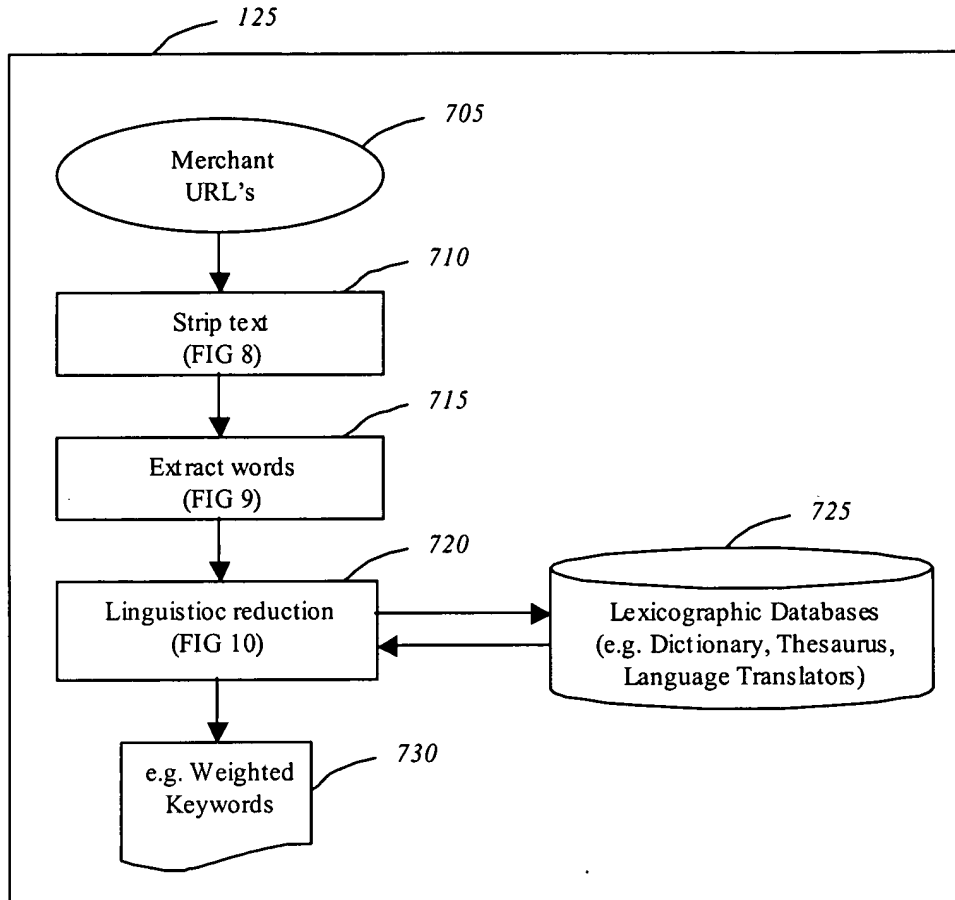
10032843-122704

Figure 6 – Validation Details



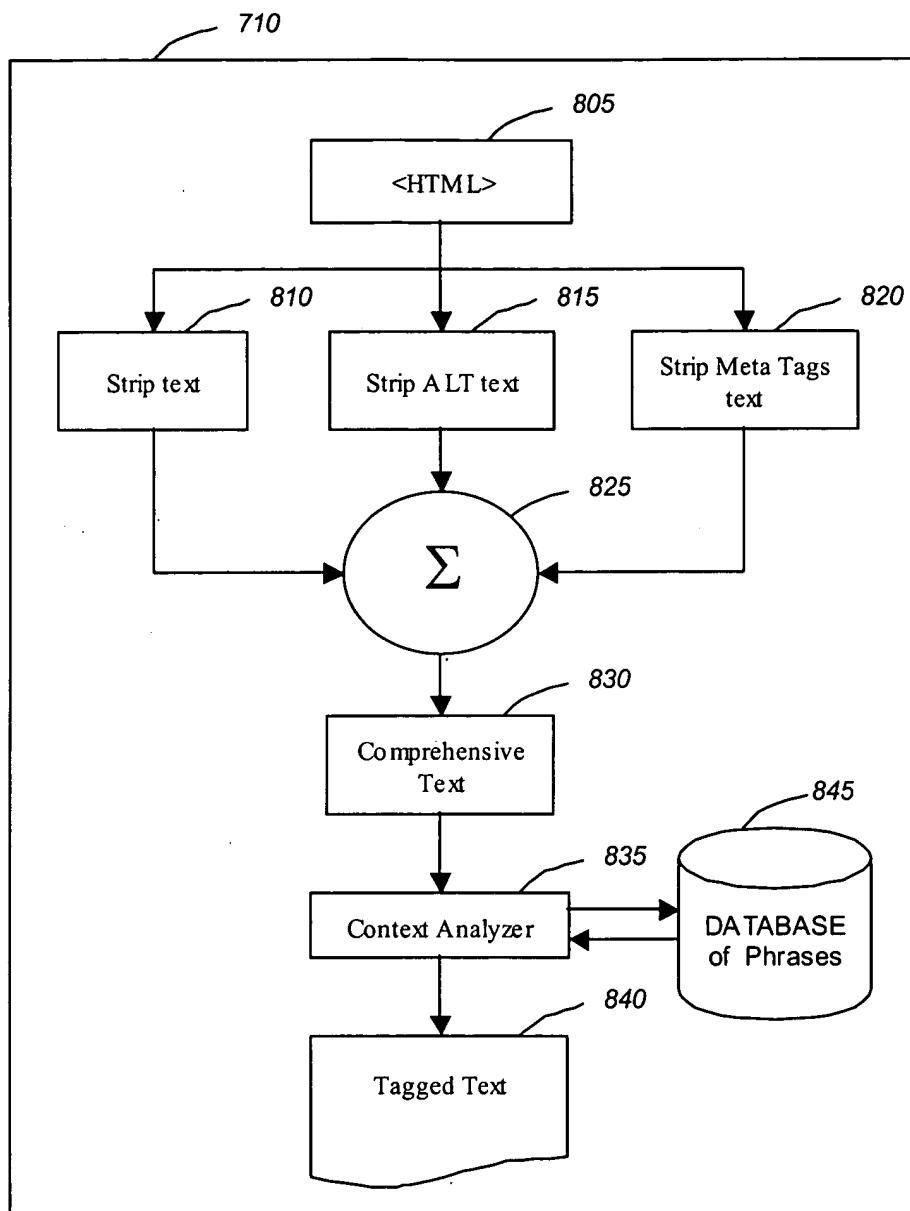
10032843.122701

Figure 7 – Analyzer High Level



10032843 122701
10/22/21 14:22:00

Figure 8 – Strip Text



10032243-122701

Figure 9 – Extract Words

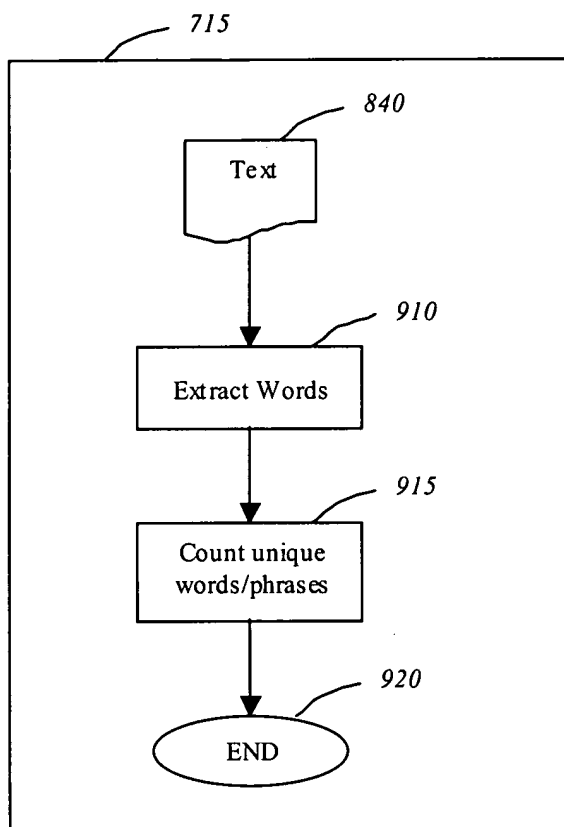
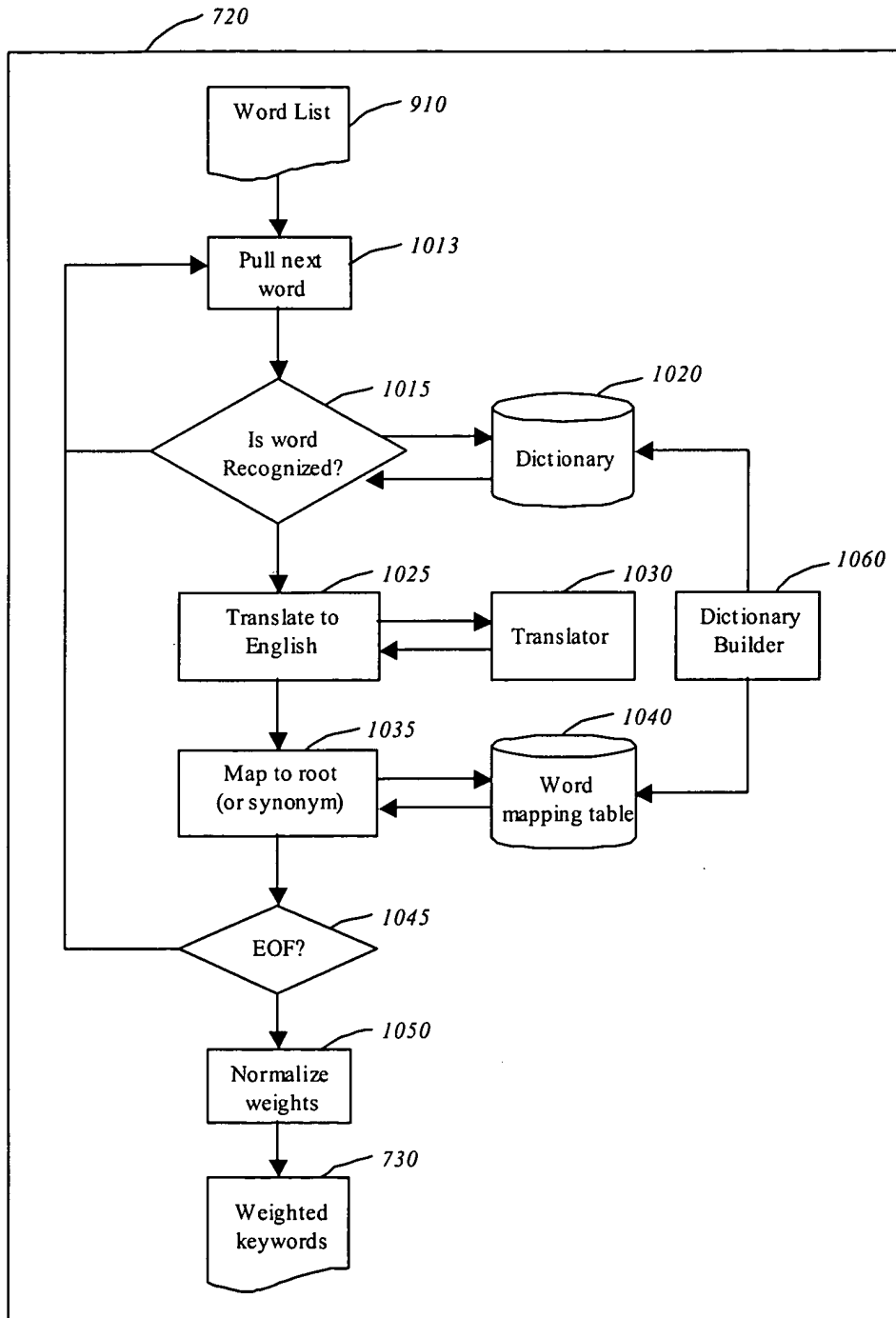


Figure 10 – Linguistic Reduction



10032343-122701

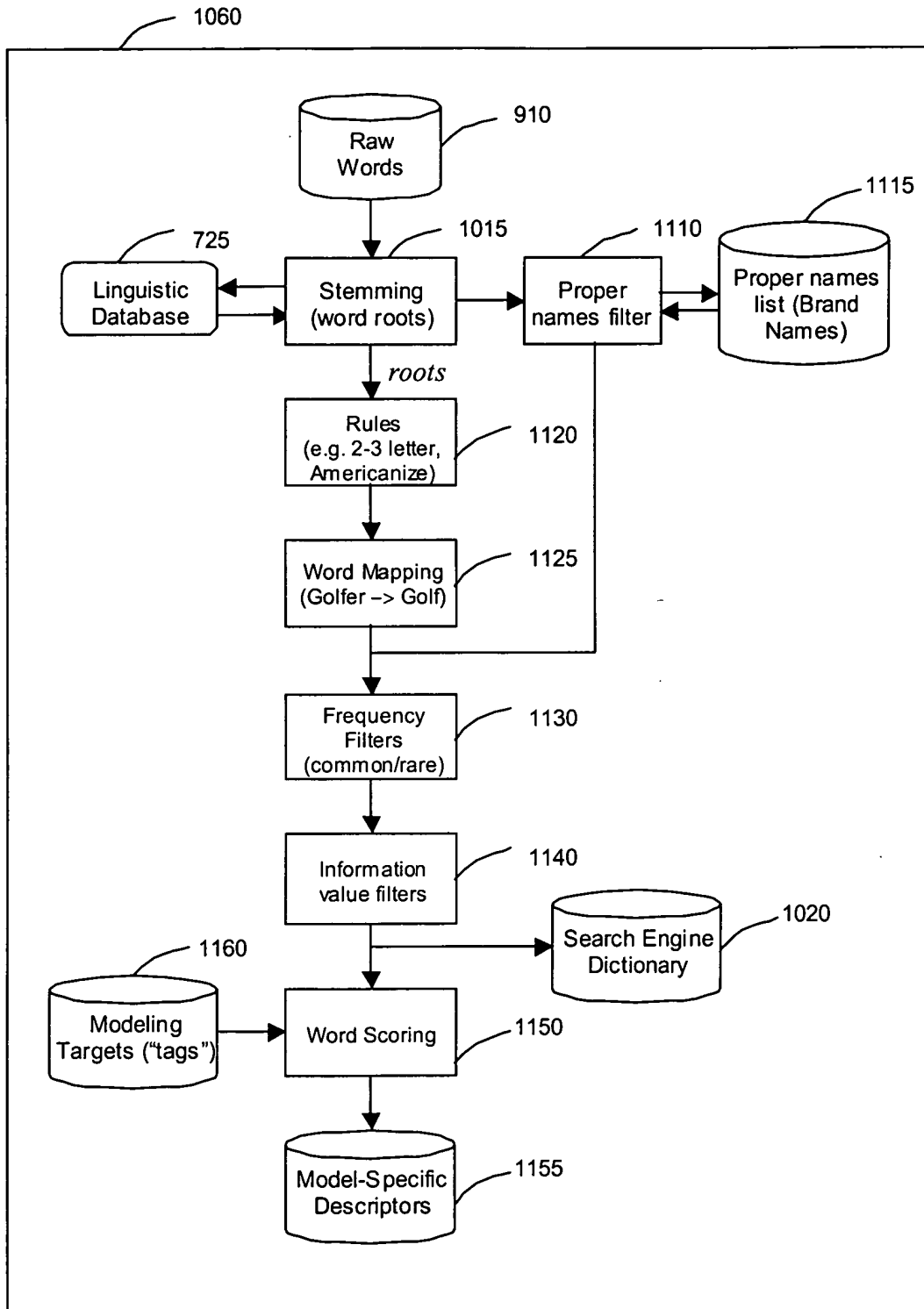


Figure 11 – Dictionary Builder

Figure 12 - Modeling

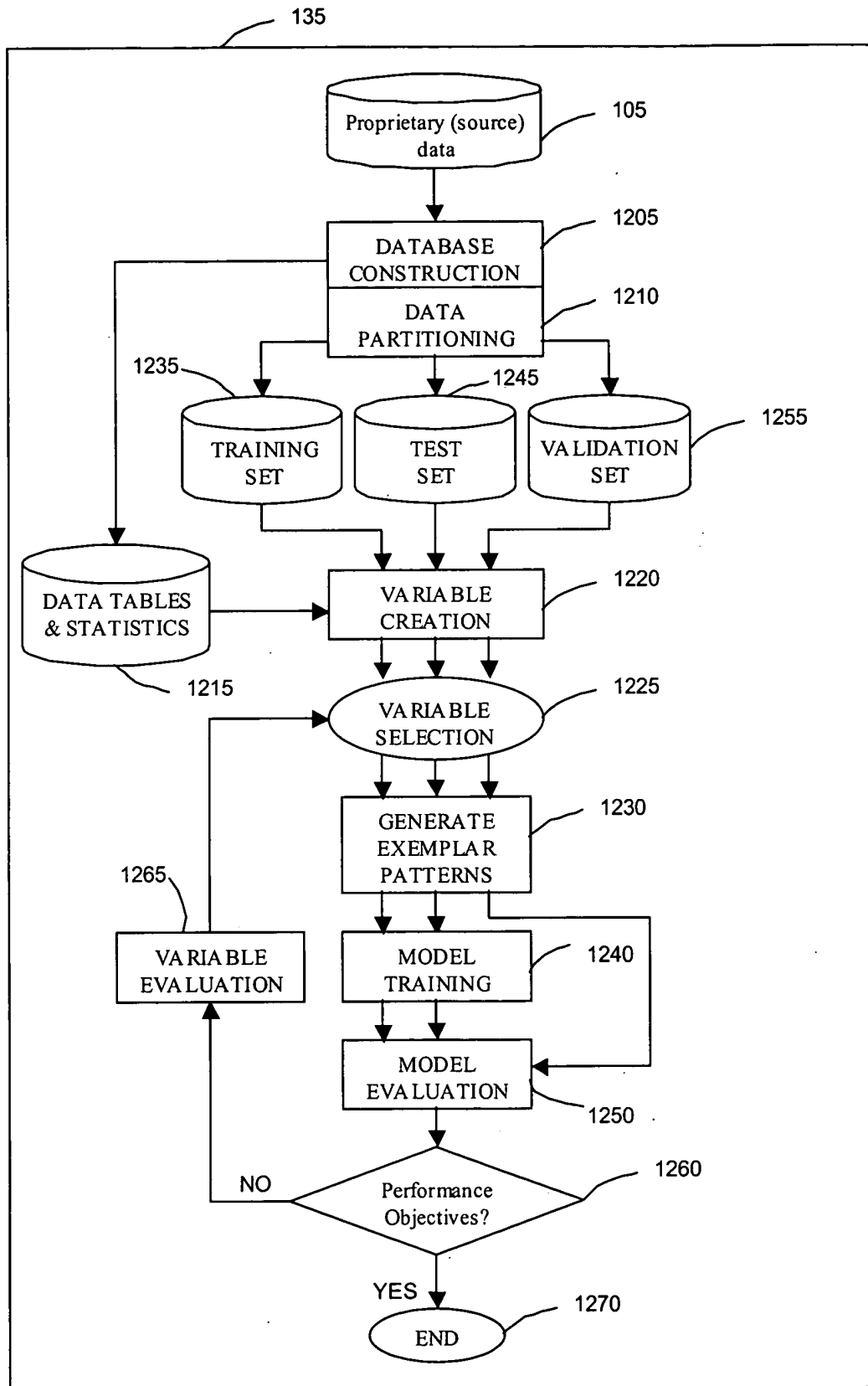
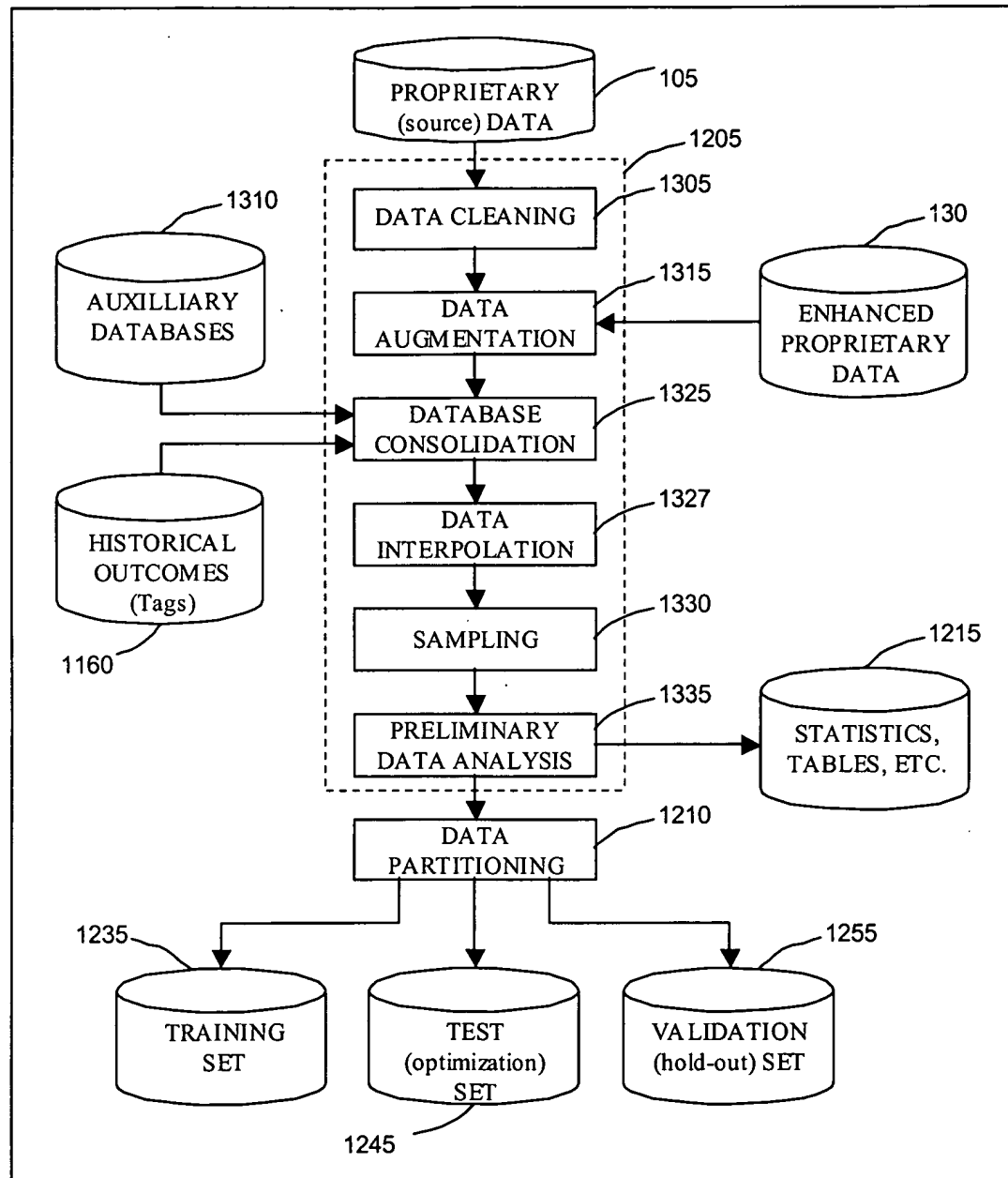
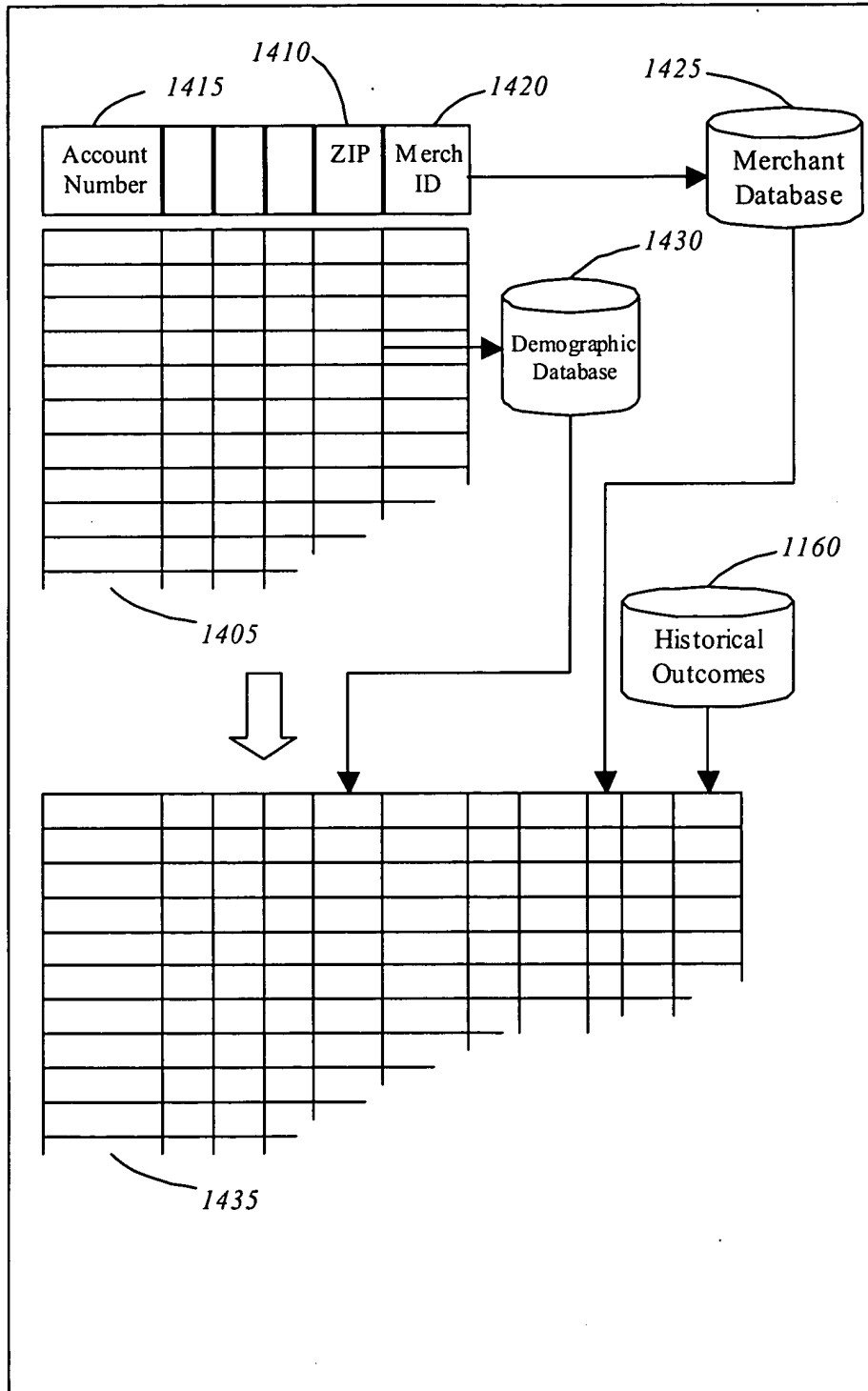


Figure 13 – Modeling Database Construction



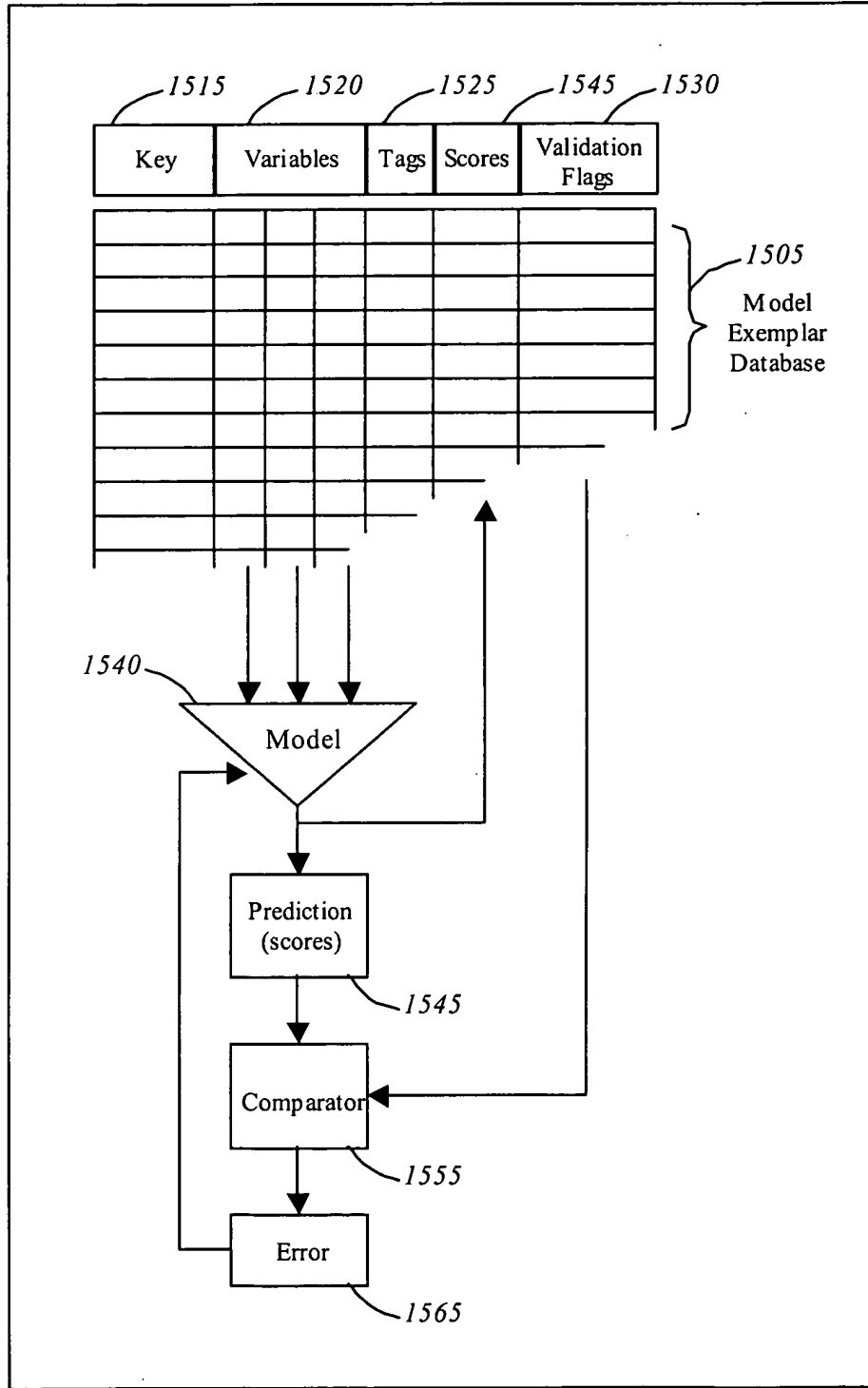
10/22/2017 10:22:01

Figure 14 – Data Augmentation Example



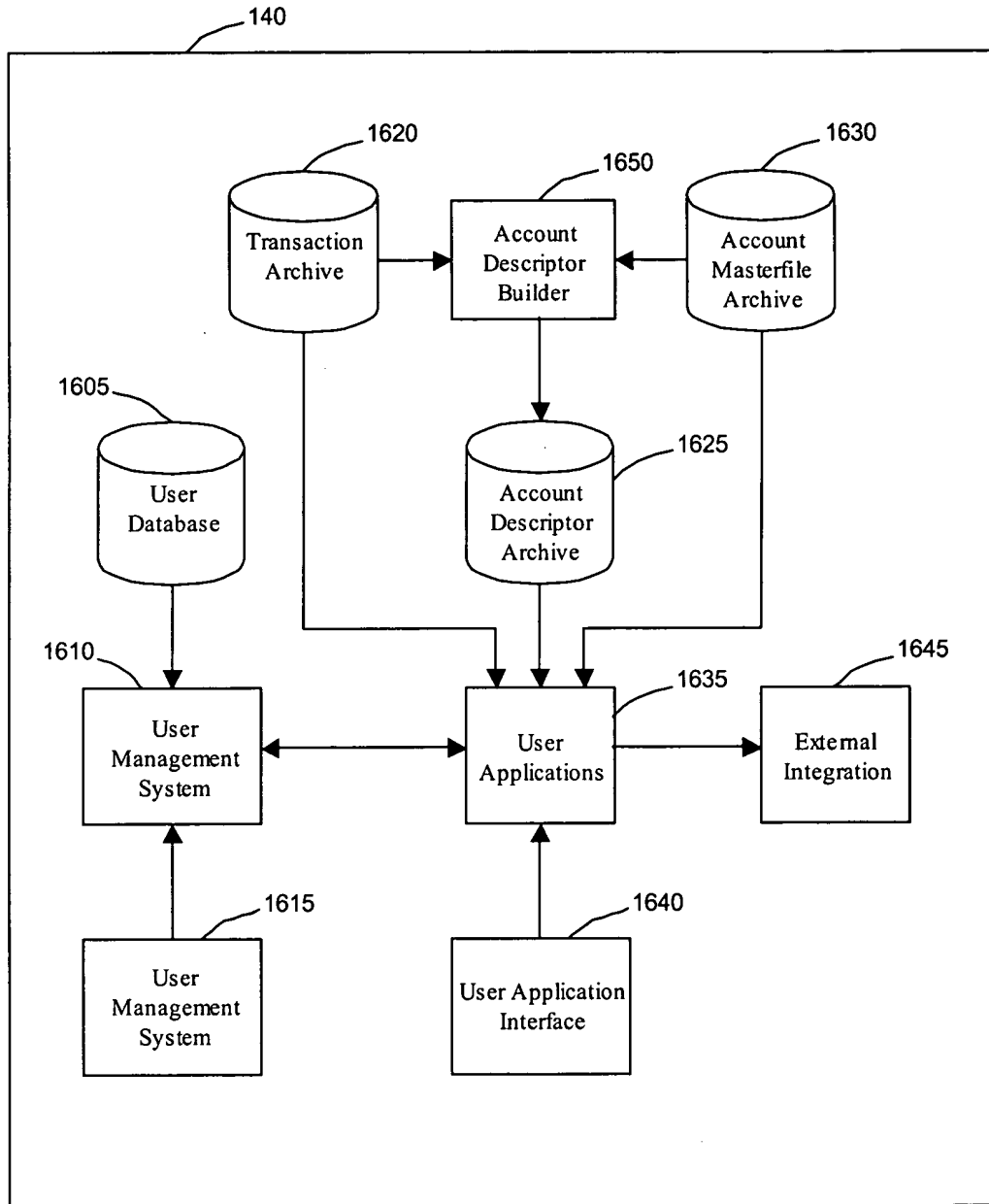
100384-12201
FOIA b 7 - D

Figure 15 – Model Training



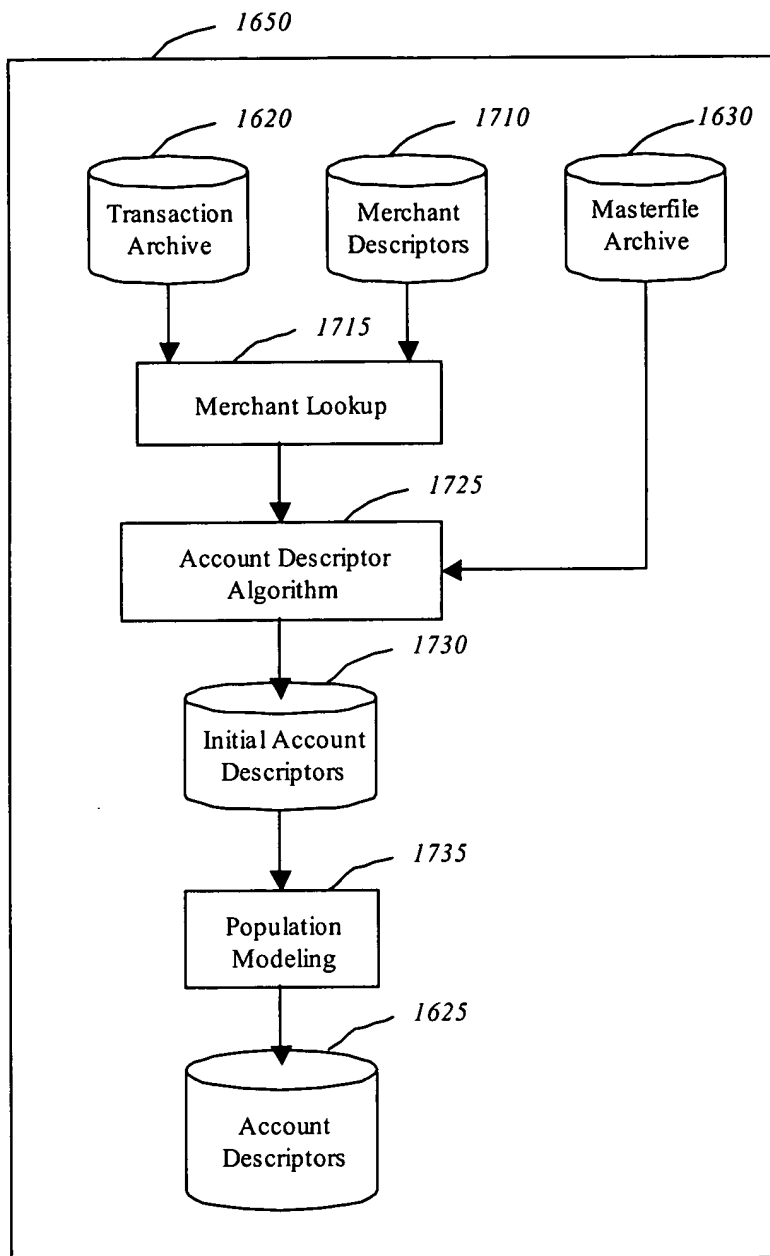
10032843-122701

Figure 16 – Data Mining Main



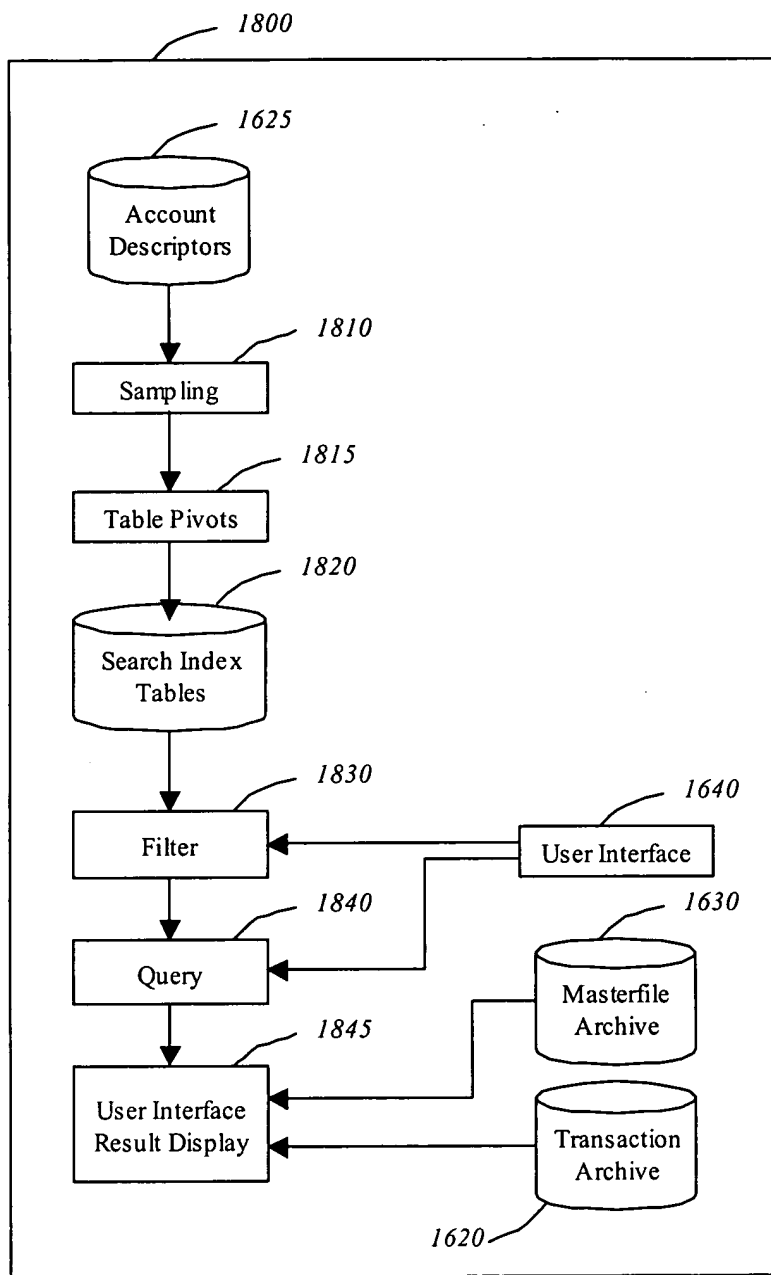
10032843 122701

Figure 17 Account Descriptor Builder



10032843-122701

Figure 18 – Search Engine



10032843 "122704
FOC22T" E482E00T

Figure 19 – Search Engine Interface

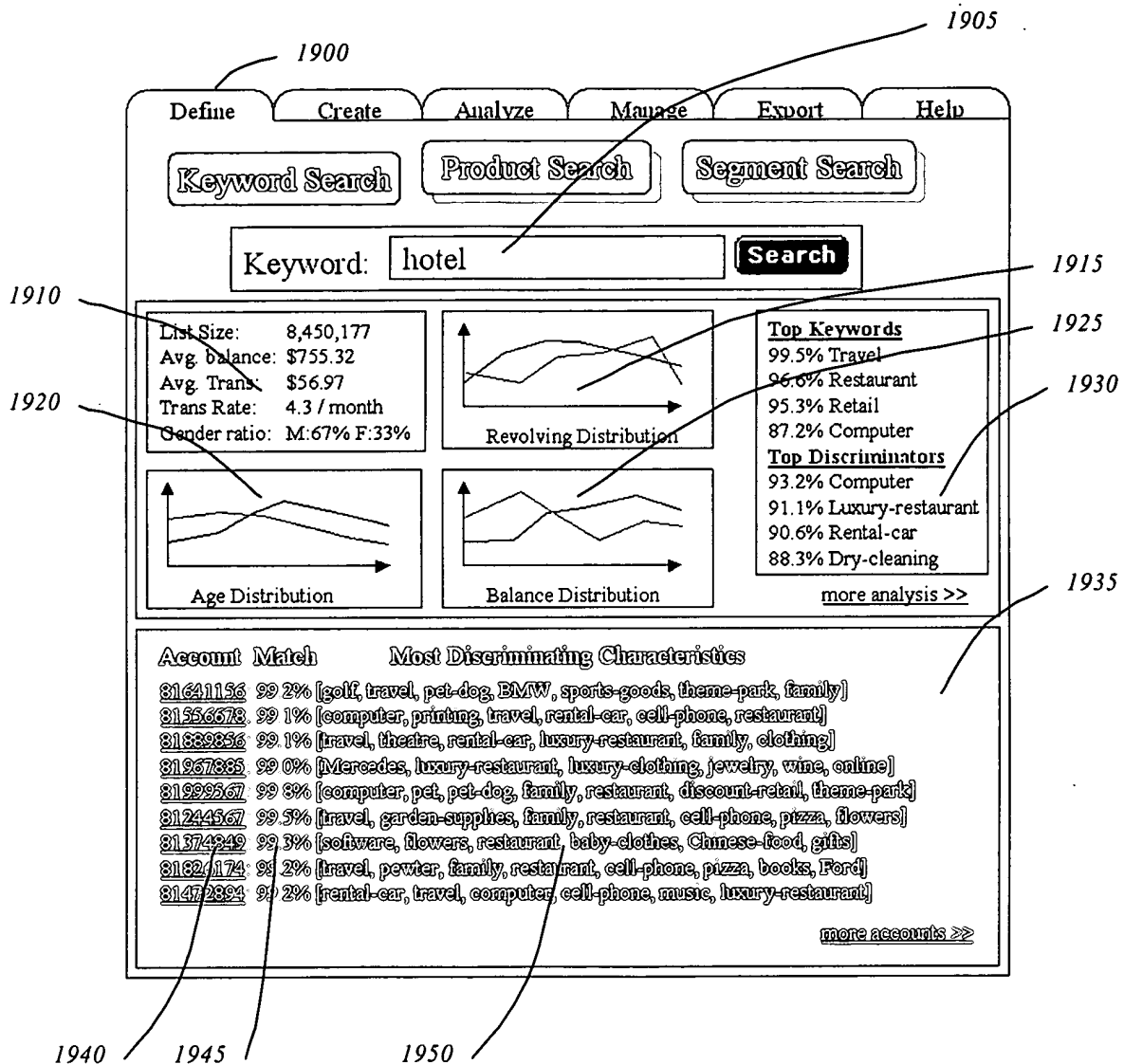


Figure 20 – Segment Description

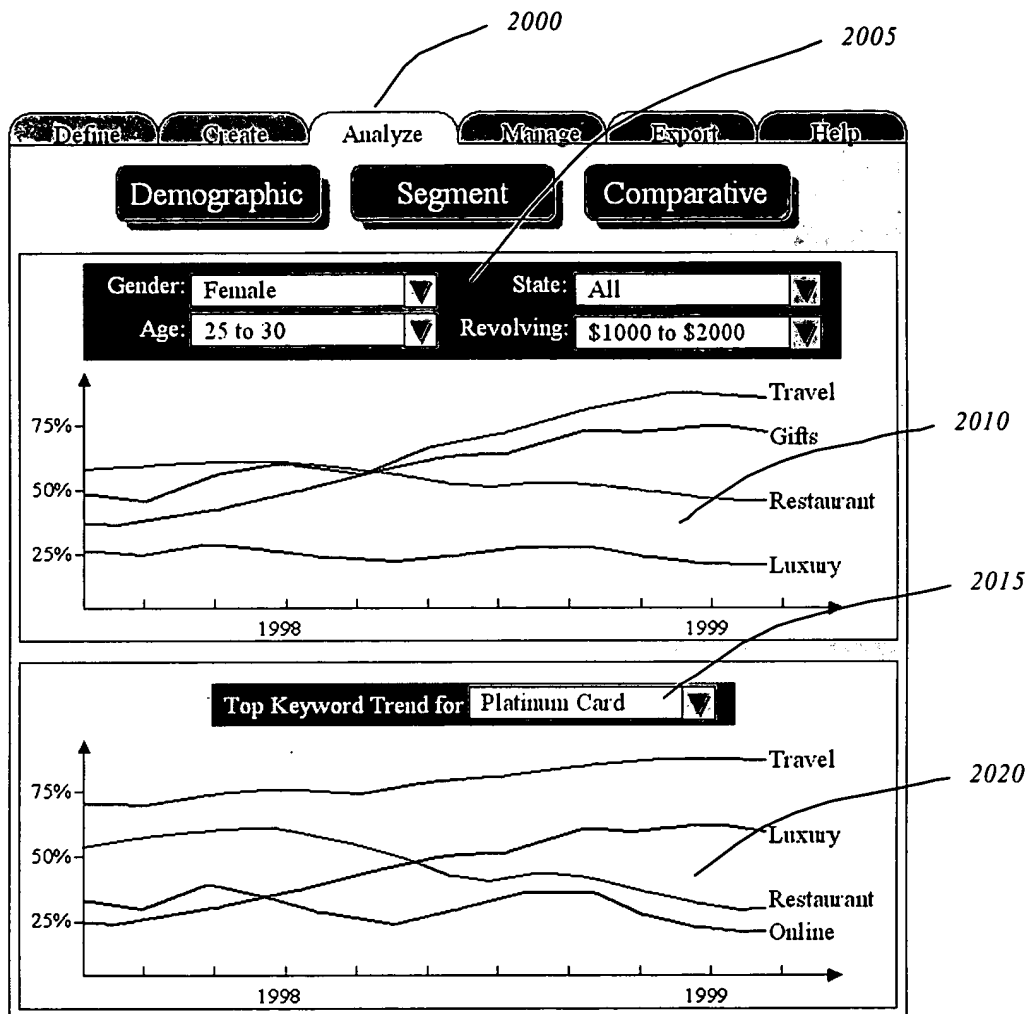
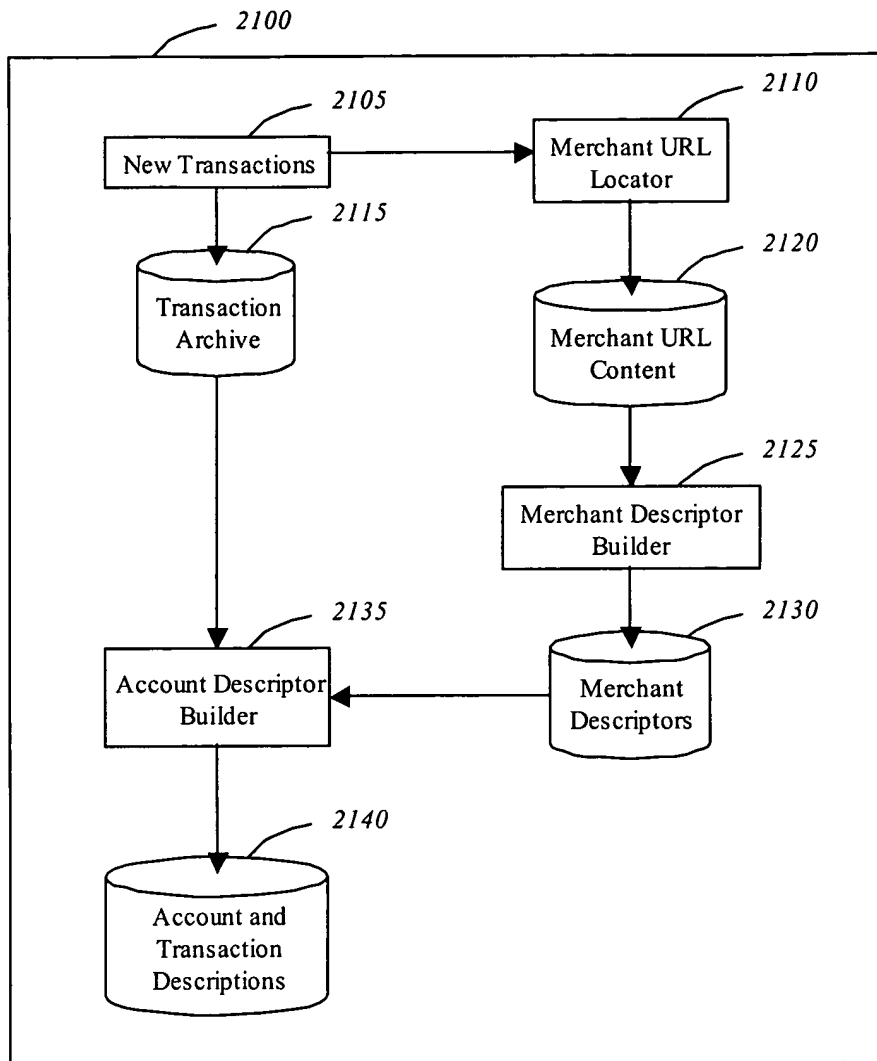
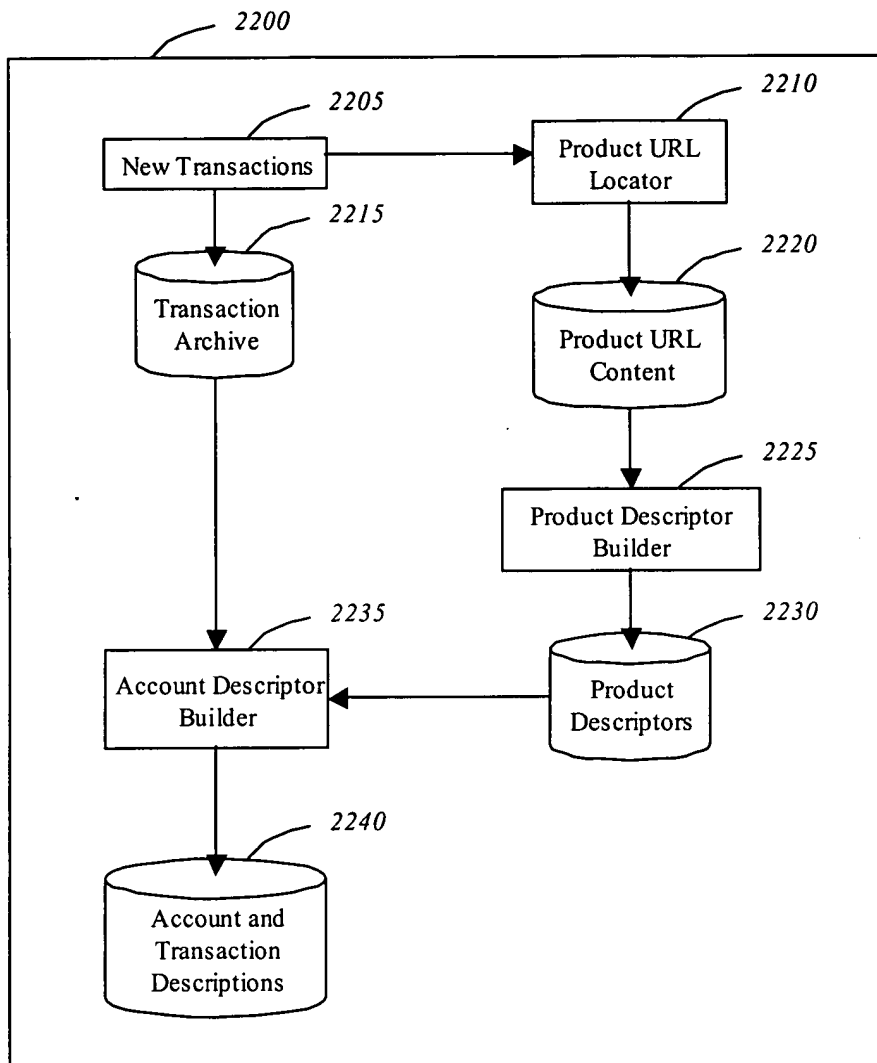


Figure 21 – CC Example



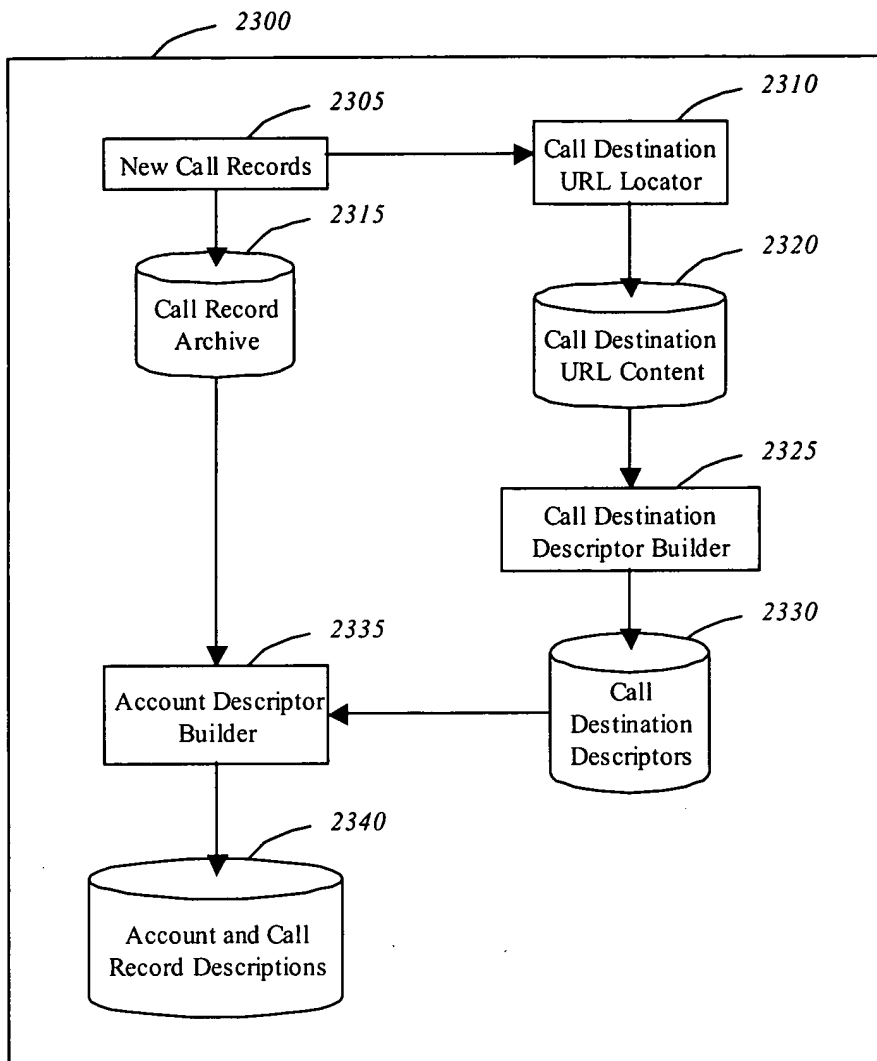
1002843 "1002843" E482E00F

Figure 22 – SKU Example



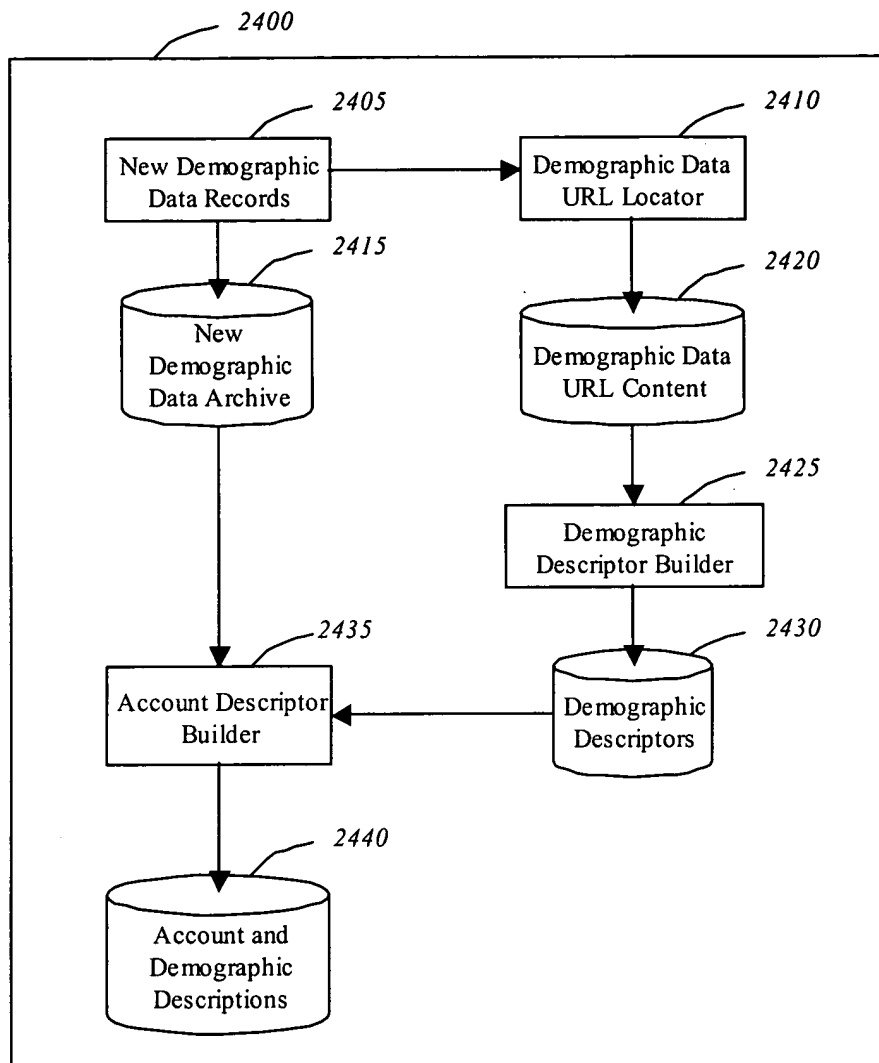
Attorney Docket "E482E001"

Figure 23 – Telephone Example



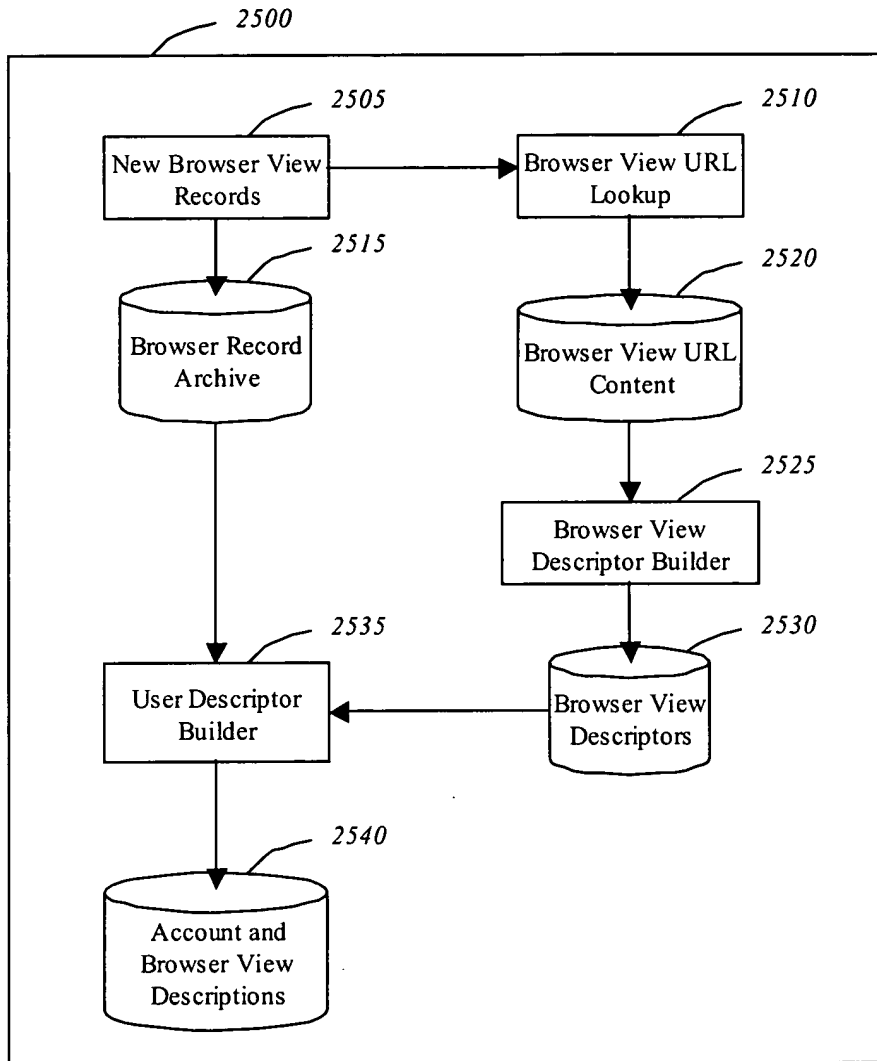
10032843 122701

Figure 24 – Demographic Data



10032843-122701
T0222T E482E00T

Figure 25 – Browser Transactions



1003241220
"E482E00T"